Habitat-Matterport 3D Dataset (HM3D): 1000 Large-scale 3D Environments for Embodied AI

Year: 2021 | Citations: 525 | Authors: Santhosh K. Ramakrishnan, Aaron Gokaslan, Erik Wijmans, Oleksandr Maksymets, Alexa

Abstract

We present the Habitat-Matterport 3D (HM3D) dataset. HM3D is a large-scale dataset of 1,000 building-scale 3D reconstructions from a diverse set of real-world locations. Each scene in the dataset consists of a textured 3D mesh reconstruction of interiors such as multi-floor residences, stores, and other private indoor spaces. HM3D surpasses existing datasets available for academic research in terms of physical scale, completeness of the reconstruction, and visual fidelity. HM3D contains 112.5k m^2 of navigable space, which is 1.4 - 3.7x larger than other building-scale datasets such as MP3D and Gibson. When compared to existing photorealistic 3D datasets such as Replica, MP3D, Gibson, and ScanNet, images rendered from HM3D have 20 - 85% higher visual fidelity w.r.t. counterpart images captured with real cameras, and HM3D meshes have 34 - 91% fewer artifacts due to incomplete surface reconstruction. The increased scale, fidelity, and diversity of HM3D directly impacts the performance of embodied AI agents trained using it. In fact, we find that HM3D is `pareto optimal' in the following sense -- agents trained to perform PointGoal navigation on HM3D achieve the highest performance regardless of whether they are evaluated on HM3D, Gibson, or MP3D. No similar claim can be made about training on other datasets. HM3D-trained PointNav agents achieve 100% performance on Gibson-test dataset, suggesting that it might be time to retire that episode dataset.